National Library of Canada
Canadian Theses Service
Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Services des thèses canadiennes

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ MICROFILMÉE TELLE QUE NOUS L'AVONS REÇUE

NL-339 (r.86/06)

Canada

THE UNIVERSITY OF ALBERTA

Comparison of the Wechsler Adult Intelligence Scale-Revised and the Multidimensional Aptitude Battery

by

Kathleen A. Ackerman

©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

AND RESEARCH IN PARTIAL FULFILMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF EDUCATION

IN

COUNSELLING PSYCHOLOGY

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA
FALL 1986

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-32334-5

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Kathleen A. Ackerman

TITLE OF THESIS : Comparison of the Wechsler Adult Intelligence Scale-Revised and the Multidimensional Aptitude Battery.

DEGREE : Master of Education

YEAR THIS DEGREE GRANTED : 1986

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

Kathleen A. Ackerman

(Student's signature)

Permanent address :

52026 Range Road 232

Sherwood Park, Alberta

T8B 1B3

Date : Oct 10 1986

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "Comparison of the Wechsler Adult Intelligence Scale-Revised and the Multidimensional Aptitude Battery" submitted by Kathleen A. Ackerman in partial fullfillment of the reqirements for the degree of Master of Education in (Counselling Psychology).

Supervisor

Date : September 19 1986

Abstract

The purpose of this study was to determine the extent to which the Multidimensional Aptitude Battery (M.A.B.) is a suitable alternative for the Wechsler Adult Intelligence Scale-Revised (W.A.I.S.-R.). The M.A.B. is available in computerized form as well as standard paper and pencil form. The study was designed to assess the comparability of these two forms in addition to their comparabilty to the W.A.I.S.-R.. A total of 40 subjects who volunteered for the study through the University of Alberta Education Clinical Services was administered the W.A.I.S.-R.. Twenty of these subjects were administered the computerized M.A.B. on an I.B.M. microcomputer, and twenty were administered the M.A.B. in the usual paper and pencil format. Half of the subjects received the W.A.I.S.-R. first and half received the M.A.B. first.

Results from the tests were first analyzed using a Hotelling T^2 test to determine whether there was a practice, or sequence, effect depending on which test, W.A.I.S.-R. or M.A.B., was administered first. The results of the analysis indicated that there was no practice effect. Secondly, a Hotelling T^2 test was performed comparing the scores from the paper and pencil M.A.B. and the scores from the computer M.A.B.. The results from this analysis indicated that there were no significant differences between the scores from the two tests. It was therefore decided to collapse the scores from the paper and pencil and computer test for further analysis. Thirdly, a correlated t test was performed to determine whether there were significant differences between scores from the W.A.I.S.-R. and the M.A.B.. Statistically significant differences were found on eight of the ten comparable subscales, the Verbal I.Q. scale and the Full I.Q. scale. Statistically significant differences were not found on the Performance I.Q. scale. Finally, in order to examine the degree of relationship between scores from the W.A.I.S.-R. and the M.A.B., Pearson Product Moment correlations were computed. Correlations for comparable subtests ranged from 0.20 for the Comprehension subtest to 0.74 for the Vocabulary subtest. Correlation coefficients for the Verbal I.Q., Performance I.Q. and Full Scale I.Q. were 0.76, 0.81, and 0.82 respectively.

As a preliminary inquiry into the question of whether comparable subtests of the W.A.I.S.-R. and the M.A.B. were measuring the same cognitive skills, after completing the two tests, a sample of 15 subjects was interviewed concerning the cognitive strategies they had used. Subjects were also asked to comment on which test they preferred. The subjects experienced a great deal of difficulty in describing the strategies they had used. The majority of subjects reported using the same strategies on comparable W.A.I.S.-R. and M.A.B. subtests, however, the strategies which they described using on each test were actually quite different. Further research using behavioral observations, in depth interviews and factor analysis is needed before conclusions concerning whether the two tests measure the same cognitive skills can be reached. In terms of test preference, 11 of the 15 subjects surveyed preferred the W.A.I.S.-R. to the M.A.B.. Those 4 subjects who preferred the M.A.B. were administered the test on the computer.

It was concluded from the study that the M.A.B. is a suitable alternative to the W.A.I.S.-R. when used for obtaining the overall level of cognitive functioning from the Verbal, Performance and Full Scale I.Q. scores. Statistically significant differences between scores on 8 out of the comparable 10 subtests of the M.A.B. and W.A.I.S.-R. suggest that the two tests do not produce equivilant subtest profiles. The question of whether the W.A.I.S.-R. and M.A.B. scales are measuring the same cognitive skills needs further research. Low correlations between some of the comparable M.A.B. and W.A.I.S.-R. subtests suggest clinical interpretations of the M.A.B. scales using W.A.I.S.-R. formats should be done with caution. Absence of a practice effect when the two tests are administered back to back suggests the M.A.B. may be used as a retest instrument for the W.A.I.S.-R.. The computerized M.A.B. was well accepted by subjects, was easy to understand, and took little time to administer, making it a potential time and cost saver for psychologists.

Table of Contents

Chapter		Page
I	Introduction	1
	Introduction to the Problem	1
	Statement of the Problem	3
	Limitations	4
II	Review of the Literature	6
	The Wechsler Adult Intellignce Scale-Revised	6
	The Multidimensional Aptitude Battery	9
	Computer-Assisted Testing	12
III	Methodology	18
	Subjects	18
	Procedure	18
	Description of Test Instruments	19
	Research Questions	22
	Analysis	23
IV	Results	25
	Description of the Sample	25
	Sequence Effect of Test Taking	25
	Comparablility of the Different Forms of the M.A.B.	25
	Correlated t Test Results	26
	Correlations	26
	Cognitive Strategies and Test Preference	29
V	Summary, Discussion and Conclusions	33
	Summary	33
	Discussion	35
	Conclusions	41
	Suggestions for Further Research	42
References		44
Appendix I		47

List of Tables

Table		Page
I	Means, Standard Deviations, t-values, and Probabilities of Differences Between Means of Subtest Scores and I.Q. Scores on the W.A.I.S.-R. and the M.A.B.	27
II	Correlations Between the W.A.I.S.-R. and the M.A.B.	28
III	Comparison of Correlation Coefficients Between the M.A.B. and the W.A.I.S.-R. and the W.A.I.S. and W.A.I.S.-R.	39

Chapter 1

Introduction

Introduction to the Problem

Intelligence testing plays a major role in personality assessment, educational assessment and vocational counselling and makes up a sizable portion of the psychologists time devoted to assessment. In addition, intelligence testing takes up a sizable portion of most psychologists' professional time and every year hundreds of hours and thousands of dollars are spent on intellectual assessments. Many psychologists and psychological corporations do exclusively intellectual assessments, yet despite the hours and money devoted to this area of psychology the demand for intelligence testing far outweighs the availability of services.

The most widely used test of adult intelligence in North America is the Wechsler Adult Intelligence Scale - Revised (W.A.I.S.-R.)(Lindemann and Matarazzo, 1984; Sternberg, 1986). Published by Wechsler in 1981, the W.A.I.S.-R. is an individually administered test designed for adults aged 16 years and older. It is a revision of the original Wechsler Adult Intelligence Scale (W.A.I.S.) which was published in 1955, with updated content and new norms based on the responses and scores obtained from contemporaneous samples of the population. The W.A.I.S.-R. is comprised of 11 subscales, six of which constitute the Verbal Scale and five of which constitute the Performance Scale. Administration of the W.A.I.S.-R yields individual subscale scores, a Verbal I.Q. based on the six Verbal Subscale scores, a Perfomance I.Q. based on the five Performance Scores and a Full Scale I.Q based on overall performance.

Interest in the W.A.I.S.-R. and the extent of its use is attested to by more than 3,000 publications on its scales to date (Matarazzo, 1985). Matarazzo (1985) described the W.A.I.S.-R. as the best standardized test for individual administration, suggesting that no other current intelligence test is as reliable, valid, or clinically useful for assessing the measurable aspects of adult intelligence. Yet despite its excellent reviews, the W.A.I.S.-R. lends itself to some

difficulties. Administration requires the expertise of trained examiners, and administration time ranges from 60 to 90 minutes, with additional time required for scoring and interpretation. As such it is costly to administer, with a limited number of qualified personnel available to administer it. Psychologists will often forego its use for the sake of other less worthy but more time efficient tests or will opt to administer a shortened but less valid and reliable form of the W.A.I.S.-R.. Because of the high demand for W.A.I.S.-R. assessments and the lack of qualified professionals to administer it clients are often put on a long waiting list. This results in a delay in clinical treatment where such diagnostic assessments are necessary for generating treatment recommendations. The W.A.I.S.-R. is often tedious to administer and the time taken to give the test could be used to administer other important diagnostic tests.

It thus becomes necessary to examine suitable alternatives to the W.A.I.S.-R.. Finding a comparable substitute would not only prove valuable in overcoming time and cost limitations, but would also provide a retest alternative. For example, it is often necessary for psychologists to retest clients in cases where the first assessment is determined to be unreliable, or where measures of change or improvement are desired. Retesting with the same test yields less valid and reliable scores due to practice effects; a suitable alternative is therefore required. This research examines the Multidimensional Aptitude Battery as one such alternative.

The Multidimensional Aptitude Battery (M.A.B.) is a recently published group administered paper and pencil test designed to measure the same intellectual factors as the W.A.I.S.-R. (Jackson 1984). The M.A.B is comprised of10 subtests, five of which make up the Verbal Scale and five of which make up the Performance Scale. Similar to the W.A.I.S.-R., the M.A.B. yields a Verbal I.Q., a Performance I.Q. and a Full Scale I.Q. Studies by its author state that scores on this test are highly predictive of W.A.I.S.-R. scores ($r=.91$ for full scale and $r=.44$-$.89$ for subscale scores). In a recent study, Hirsch (1986), found that group administered scores on the M.A.B. with adults returning to school were quite comparable on the Verbal section of the test, whether it was given on a computer or in a regular group administration. In the Hirsch study, the verbal

section of the test was computerized on a program developed specifically for the study; the performance section of the test was not computerized due to difficulties in producing the necessary graphics.

Since the Hirsch study , Jackson has produced a computerized version of the M.A.B. including both the verbal and performance sections of the test. The computerized version of the M.A.B. has the potential to save considerable time and money for psychologists, thus allowing more extensive and thorough assessments. It has a built-in timing factor and is self explanatory, unlike the paper and pencil version, which requires that the psychologist be present to time each seven minute section as well as read aloud instructions to clients. With the computerized M.A.B. the psychologist is free to attend to other work while the client completes the test on the computer. In addition, the computerized M.A.B. has a built-in program for scoring and can generate different types of summary reports depending on whether the test is to be used for educational or clinical purposes. This feature saves additional valuable time and makes results available immediately.

To date no comparative studies have been done with the computerized version of the M.A.B. to determine if it is equivalent to the paper and pencil M.A.B or the W.A.I.S.-R., and there is no normative data available on the computerized version. As well, the question of the validity of the computerized M.A.B. in general has not been addressed. Consistent with the findings of Hirsch (1986), Jackson has been unable to produce the graphics needed to completely administer the performance section of the M.A.B. on the computer. Clients must refer to the performance test booklet for each subtest and then enter their answers on the computer. It remains to be shown whether this is an acceptable way to administer the test.

Statement of the Problem

This study was designed to assess the extent to which the M.A.B. is a suitable alternative for the W.A.I.S.-R. More specifically, the following questions were addressed:

1. Is there a practice effect on taking both the W.A.I.S.-R. and the M.A.B.?
2. Are there any differences between the scores on the paper and pencil M.A.B. and the scores on the computerized M.A.B.?
3. Are there any differences between the Verbal, Performance, and Full Scale I.Q. scores on the W.A.I.S.-R. and the Verbal, Performance and Full Scale I.Q. scores on the M.A.B.?
4. Are there any differences between the subscale scores on the W.A.I.S.-R. and the subscale scores on the M.A.B.?
5. What is the relationship between the W.A.I.S.-R Verbal, Performance and Full Scale I.Q. scores and the M.A.B. Verbal, Performance and Full Scale I.Q. scores?
6. What is the relationship between the W.A.I.S.-R. subtest scores and the M.A.B. subtest scores?

This research will also serve as a preliminary enquiry for two other questions:

1. Are the corresponding subtests on the W.A.I.S.-R. and the M.A.B. measuring the same cognitive functions such that subjects report using comparative cognitive strategies in responding to equivalent subtests for both the W.A.I.S.-R. and the M.A.B.?
2. Are there qualitative differences between the tests, such that a preference exists for one over the other? The above questions, numbered 1 to 6, will be addressed in a statistical form in the methodology section. The secondary questions will be approached in a descriptive report format.

Limitations

Since the purpose of this study is to address the statistical comparability of the W.A.I.S.-R. and the M.A.B., only limited qualitative comparison can be made. The study does not address the question of the validity of computer generated reports. Limitations also arise when considering the nature of the sample population. Since subjects were volunteers, care should be taken in generalizing these research findings to the overall adult population.

In the following chapter the literature relating to the W.A.I.S.-R., the M.A.B., and computer-assisted testing will be reviewed.

Chapter 2

Review of the Literature

The Wechsler Adult Intelligence Scale-Revised

The Wechsler Adult Intelligence Scale-Revised (W.A.I.S.-R.) is today's most frequently used test of Adult Intelligence. The W.A.I.S.-R., designed by Wechsler in 1981, was developed for adults aged 16 to 74 years. Administration, scoring, and interpretation of the scale requires extensive post-graduate level training in its use. The test, which takes 60 to 90 minutes to complete, is designed for individual administration and examiners must adhere to rigorous standardized administration procedures as described in the manual.

Wechsler (1981, p.7) defined intelligence tests as " sets of standardized questions and tasks for assessing an individual's potential for purposeful and useful behavior." He believes that although the tests evaluate cognitive abilities, this is not their primary purpose, and "the information obtained from intelligence tests is relevant to the extent that it establishes and reflects whatever it is one defines as overall capacity for intelligent behavior"(p. 7). Wechsler's definition of intelligence tests is closely related to his conception of intelligence, defined as the "aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his or her environment. It is aggregate or global because it is composed of elements or abilities which, though not entirely independent, are qualitatively differentiable" (p.7). Wechsler's belief that intelligence involves separate abilities organized together is reflected in the construction of the W.A.I.S.-R.. The test is composed of 11 subtests, 6 of which form the Verbal Scale and 5 the Performance Scale. Administration of the test yields a Verbal I.Q. score, a Performance I.Q. score, and a Full Scale I.Q. score, or overall general ability score.

The W.A.I.S.-R. extends the line of test development that began with the Wechsler Belvue Intelligence Scale-Form I published in 1939, and continued with

its revision to the Wechsler Adult Intelligence Scale (W.A.I.S.) which was published in 1955. The primary objective of the 1939 Wechsler Scale was to provide an intelligence test that was suitable for adults (Anastasi, 1982). Previous intelligence tests had been designed primarily for school children and had been adapted for adults by adding more difficult items of the same kind.

In 1981 Wechsler published a revised version of the 1955 W.A.I.S., with updated content and new norms based on responses and scores obtained from contemporary samples of the population. Wechsler (1981) stated that his primary reason for revision and restandardization of the W.A.I.S. was "to insure its continued effectiveness as a basic test of intelligence and as a valid diagnostic tool and research instrument"(p.1). The W.A.I.S.-R contains a number of changes in content but Wechsler states that his views on the nature and intent of an intelligence test and what it measures have not undergone any marked changes in recent years. Standardization of the W.A.I.S.-R. is based on administration of the full scale to 1880 adult subjects selected according to U.S. census data and tested between 1976 and 1980. A stratified sampling plan was adopted to ensure that representative portions of adults along the variables of age, sex, race, occupation, education and urban-rural residence would be included.

Reliability coefficients for the W.A.I.S.-R. were computed by Wechsler within each of the nine age groups for each of the 11 subtests, as well as for Verbal, Performance and Full Scale I.Q.s. Corrected split half reliability coefficients for the Full Scale I.Q. ranged from .96 to .98; for the Verbal I.Q. they ranged from .95 to .97; and for the Performance I.Q. .88 to .94. The individual subtests had lower internal consistency reliabilities, ranging from .52 for Object Assembly at age 16 to17 to .96 for Vocabulary at 6 of the 9 age levels. Test- retest reliability coefficients for the subscales ranged from .69 to .91 for ages 25 to 34 and from .67 to .94 for ages 45 to 54. For the Verbal, Performance and Full Scale I.Q.s the test-retest reliability coefficients were .94, .89 and .95 respectively for ages 25 to 34 and .97, .90 and .96 respectively for ages 45 to 54.

The W.A.I.S.-R. manual itself contains no validity data. Wechsler (1981) stated "that a body of evidence, both rational and empirical, attests to the validity of the Wechsler adult scale as a measure of global intelligence"(p.14). Indeed, the

body of literature dealing with the validity of the Adult Wechsler Scales is vast and cannot be enumerated here. However, several aspects of validity are discussed in Wechsler's 1958 book. Wechsler (Chapter 5) argued that the psychological functions tapped by each of the 11 subtests fit his definition of intelligence, that similar tests have been successfully employed in previously developed intelligence scales, and that such tests have proven their worth in clinical experience.

The W.A.I.S.-R. is touted as the most reliable and valid test of adult intelligence in existence today (Matarazzo, 1985). Jackson (1984) speaks highly of the W.A.I.S.-R., reporting that "the Wechsler scales have become the standard against which other tests of intelligence have been appraised -- a new adult measure of general intellectual ability could hardly be expected to be acceptable if its correlation with the W.A.I.S.-R. was low" (p.5). Jackson believes the Wechsler scales have been successful for a number of sound reasons: (1) they incorporate a diversity of tasks including not only verbal and school learned content, but performance and practical skills as well; (2) they contain content appropriate for adults and adolescents; (3) they reflect fresh concepts of the nature of intelligence in which psychotic processes, neurological damage, or emotional disturbances might affect performance; (4) they have been very carefully normed and standardized, using United States census data; (5) they have stimulated a substantial body of research -- over 3000 research articles, monographs and books based on the tests have been published; (6) they are of high technical psychometric quality, as evidenced by very high levels of internal consistency reliability for the full scale W.A.I.S. and W.A.I.S.-R., ranging to above .95; and (7) substantial validity data have been published over the years, ranging from predictions of academic grades to studies of institutional release rate for persons classified as mentally retarded and their subsequent work adjustment (p.5). Jackson feels that the only serious drawback to the W.A.I.S.-R. is that it requires individual administration and scoring by a specially-trained professional.

The W.A.I.S.-R. has a wide range of applications. It is frequently employed as a screening instrument and is especially prevalent in the testing of normal adults and adolescents for educational and occupational counselling, personnel

selection, and similar purposes (Anastasi, 1982). Another common use is to be found in clinical testing , especially in the identification and classification of the mentally handicapped. In the clinical area it is also used as an aid in the assessment of cerebral pathology and for detecting deficits in the specific cognitive functions of attention, memory, perception, thinking and judgement (Frank, 1983). The W.A.I.S.-R. is routinely administered in psychiatric hospitals and mental health institutions for assessing general intellectual level and is often used in the assessment of schizophrenia and other psychiatric and emotional disorders.

The Multidimensional Aptitude Battery

Published by Jackson in 1984, the Multidimensional Aptitude Battery (M.A.B.) is a paper and pencil test designed to provide " a convenient objectively-scorable measure of general aptitude or intelligence in the form of a profile containing five verbal and five performance subtest scores" (p.5). The M.A.B. can be administered in one hour and 40 minutes and yields a Verbal I.Q. score, a Performance I.Q. score and a Full Scale I.Q. score. Lengthy training is not required -- administration procedures for the test can be mastered in 10 to 15 minutes-- and the test can easily be administered by nonprofessionals such as teachers, secretaries, and psychological assistants. The M.A.B. uses a five choice multiple choice format and can be administered to groups of up to 25 by a single tester. The M.A.B. is patterned after the Wechsler Adult Intelligence Scale-Revised, and Jackson claims that it measures the same abilities as the W.A.I.S.-R. even though the two tests differ in content, administration, and response formats.

The M.A.B. was revised several times prior to publication. This process involved extensive item analysis, revisions and field testing. In item selection, particular attention was paid to identifying content that would have a fair level of generality across diverse groups. The M.A.B was standardized by equating it with the W.A.I.S.-R. Using a heterogeneous sample of 160 respondents, the equating was performed in two stages, "first at the subscale level, and second, at the summary level, separately for the Verbal Scale, Performance Scale and Full

Scale"(p.30). To evaluate the validity of the equating procedure and to obtain Canadian adolescent norms, over 5000 high school students were tested. The average I.Q. of these students computed using W.A.I.S.-R.-equated M.A.B. norms was 103, which is 3 points above that expected for the normal population. Jackson reports that "this value is very close to expectations given the fact that it is more likely that persons lower rather than higher in ability will drop out of school" (p.30).

Internal consistency reliabilities were computed separately for 230 male and 258 female adolescents ranging in age from 15 to 20. Reliability coefficients for Verbal, Performance and Full Scale ranged from .94 to .97 for different age groups. Coefficients for subscales were lower, ranging from .70 for arithmetic at age 16 to .96 for Spatial at ages 19 and 20. Test retest reliability coefficients for 52 young adult psychiatric patients for the Verbal, Performance and Full Scales were .95, .96, and .97 respectively. Coefficients for the subscales ranged from .83 for the Similarities scale to .97 for the Information scale.

To test for validity, Jackson administered the M.A.B. and the W.A.I.S.-R. to a sample of 145 individuals. The resulting correlations for the Verbal Scale, Performance Scale and Full Scale were .94, .79, and .91 respectively. Correlation coefficients for the subtests ranged from a low of .44 for Spatial/Block Design to a high of .89 for Arithmetic and Vocabulary. To date, only one other study has been done comparing the M.A.B. and W.A.I.S.-R.. Because the study involved the use of a computerized M.A.B. it will be discussed in a later section. The test is listed in the most recent edition of the Mental Measurement Yearbook but is not accompanied by any reviews.

In 1986 Jackson developed a computerized version of the M.A.B. for use with I.B.M. P.C. and P.C. compatible microcomputers. The test is identical to the paper and pencil version in content and format, the only difference being that the instructions for the test are administered via the computer rather than by a tester. The computer program has a built-in timing factor, and is completely self explanatory, allowing the examinee to take the test independently. Once the computer program has been started the tester need not be present to supervise the test. Due to difficulties in producing the graphics for the performance section

of the test, when completing this section clients must refer to the Perfomance test booklet and then answer on the computer. The computer M.A.B. also has a built in program for scoring and can generate several different types of test reports depending on whether the test is intended for educational or clinical use.

Reliability, validity, and norming data are not presently available for the computerized M.A.B. To date no comparative studies have been done with the program produced by Jackson, although one study has been done by Hirsch (1986) using the paper and pencil version and a computerized Verbal M.A.B. scale developed for the study. The study compared the results from administration of the Verbal subscales of the W.A.I.S.-R., the paper and pencil M.A.B., and the computer M.A.B. developed by Calder. Subjects were 40 adults attending the Alberta Vocational Center, which is a school for economically and educationally deprived adults who do not qualify to enter regular post secondary institutions. An Apple IIE computer was used to administer the M.A.B. to 20 subjects while the other 20 were administered the paper and pencil version. All subjects were administered the Verbal subscales of the W.A.I.S.-R. The results of the Hirsch study indicated that there were no significant differences between test scores from the computer administered M.A.B. and test scores from the paper and pencil M.A.B. Analysis for differences between the M.A.B. and the W.A.I.S.-R. scores was not performed. Correlations between the W.A.I.S.-R. and the paper and pencil M.A.B. ranged from .55 for similarities to .81 for vocabulary. The correlation coefficient for the Verbal I.Q. Scale was .83. For the W.A.I.S.-R. and the computerized M.A.B., correlation coefficients for the subscales ranged from .59 for Comprehension to .85 for Vocabulary. The coefficient for the Verbal I.Q. was .86.

According to Jackson (1984), the M.A.B. may be used in a variety of settings, including educational and career counselling, business and industry, clinics and mental health facilities, and in basic research. For adequate assessment subjects must be able to read and understand the written directions. It is recommended by Jackson that the test not be used in assessing the level of intellectual functioning of individuals suspected of mental retardation, nor for psychotic individuals whose thought processes might interfere with understanding the directions. Jackson (1984) states that his personal experience however, has been "that the majority of

psychotic patients are capable of purposefully completing the M.A.B." (p.15).

Computer Assisted Testing

The use of computers in psychological assessment has increased dramatically over the past ten years. There is a relatively large body of current literature dealing with the different aspects of computer-assisted testing. The present review will highlight some of the literature examining the benefits of computer-assisted testing, potential problems, and implications for the future.

The focus of this review will be on assessment programs which administer, score, and interpret tests "on site" and provide immediate feedback through computer analysis. These programs normally run on microcomputers such as the Apple and the IBM PC, although there are large expanded systems which run on a large minicomputer system such as the Digital Equipment Corporation VAX. In addition, there are a number of computer assisted scoring systems available for private use and there are several computer scoring services available by mail.

The principal manufacturer of software programs for individual use is a company called Psyc-Systems in Baltimore, Maryland. Brown (1984) reported that as of 1984 the Psych-Systems software consisted of the following tests: Minnesota Multiphasic Personality Inventory, Medical History Survey, Career Assessment Systems, California Psychological Inventory, Rorschach Interpretive System, Jenkins Activity Survey, Multidimensional Aptitude Battery, Self Directed Search, Vocational Preference Inventory, Visual Searching Task, Eysenck Personality Questionnaire, Adjective Checklist, Spielberger, State Trait Anxiety Scale, Beck Depression and Hopelessness Scale, Behavioral Observation Checklist, Minnesota Child Development Scale, and Personality Inventory for Children. The presentation format for most of the tests is either true/false or multiple choice, although in some cases the computer will accept free response patterns from the client. "Scoring of the instruments ranges from simple production of stanines and/or percentiles to elaborate presentations of factor scores based on multiple regression techniques" (Brown, 1984, p.458).

Brown (1984) suggested that the benefits of computer-assisted testing include the following: (a) Tests can be administered quickly without tying up

excessive amounts of professional time, (b) The administration of the tests is perfectly standardized and thus their reliability is enhanced, (c) Both children and adults appear to enjoy this type of testing more than traditional psychological evaluation, (d) Scoring and report writing are enhanced by the rapid feedback provided by computerization, (e) The ability of the computer to accurately score and interpret test data using highly sophisticated statistical techniques rivals and/or exceeds capabilities of professional psychologists (p.458).

Sampson (1983) discussed a more extensive list of potential benefits:

1. Positive client response. Sampson reported that a common criticism of computer testing is their dehumanizing effect on the client, however after reviewing the literature he concluded that overall, clients respond positively to computer tests. This finding is also supported by Burk and Normand(1986) who examined the attitudes of 217 clerical employees toward computer-administered psychological testing. Fowler(1985) suggested that there is nothing intrinsically dehumanizing about computerized testing and stated that "a client may be dealt with warmly and sensitively or coolly and impersonally regardless of the modality" (p.754). After reviewing the literature he concluded that fear of computers is a malady that affects professionals much more that their patients.

2. Cost Effectiveness. Sampson reported that in comparison to traditional testing and assessment methods, computer applications have proven to be cost effective.

3. Adaptive testing. This allows the administration of items which are most appropriate to the individual, thus eliminating unnecessary questions and saving time.

4. Generation of ancillary test data. In addition to recording answers, the computer is also capable of recording data on how items are completed. Sampson stated that "This has the potential for providing improved general instrument validity and specific scale validity for individual testing"(p.295).

5. Staff efficiency. Since the computer can handle many clerical tasks, support staff and professionals have more time available to pursue other tasks.

6. Administration and scoring efficiency. In comparison with traditional methods, the time necessary to complete computer-assisted testing and to

generate an interpretation is reduced.

7. Reduced error rates. Because items are usually presented one at a time, errors due to being out of synch between the answer sheet and test booklet are eliminated.

8. Validity of interpretive reports. Sampson reported that in comparison to clinical judgements, computer generated interpretations appear to provide equal or superior validity and reliability. However, Sampson only provided one reference to support his statement. Matarazzo(1986) is much more conservative in his evaluation of computer generated interpretive reports. He stated that

> computerized clinical psychological test interpretations offer considerable potential for improving the work of health service providers.....However, until future research establishes that such descriptions meet even the most primitive scientific tests of validation, let alone more adequate ones, it is essential that they be used only as tools by the clinician trained in their use and not as equivalents of, and thus substitutes for, professional education and training (p.14).

9. Assistance to individuals with visual, auditory, and physical limitations. Microcomputers with specialized data input and output devices have the potential to provide handicapped individuals opportunities to complete various tests with minimal staff assistance.

10. Research opportunities. The unique data collection and storage capabilities of the computer make it feasible to conduct research that was previously prohibitively difficult or impossible. In keeping with this, Brown(1984) suggested that computers can significantly improve the time frame for the development and implementation of a test such that in the future tests could be developed and standardized in one or two years rather than 5 to 7 years.

In addition to benefits, Sampson(1984) perceived that there are a number of potential problems with computerized tests. Sampson's views are supported throughout the literature and, because he has presented the most comprehensive list of problems, his main points will be presented and supporting literature will be

discussed under his headings.

1. Inadequate provision for human factors (Sampson, 1984; Hofer and Green, 1985). Sampson suggested that systems which incorporate established human factor principles into their design tend to be easier to understand and use. Several human factor problems which he lists are: misunderstanding instructions, difficulty in reading, pressing the wrong key, and long time lags between items. Sampson feels that client anxiety about their ability to operate a computer can result from poor system design.

2. Inadequate client screening (Sampson, 1984; Hofer and Green, 1985; Brown 1985). Sampson suggested that individuals in crisis may not be able to adequately respond to a computer system and that there is a need for initial screening to determine whether potential users are psychologically or intellectually capable of successfully using a computer system.

3. Confidentiality (Sampson, 1984; Sampson and Pyle, 1984). Computers magnify the potential for abuse of invasion of privacy and unauthorized access to test results.

4. Inaccurate generalized test interpretations (Sampson, 1984; Fowler,1985; Matarazzo,1986). Sampson suggested that generalized test interpretations need to be presented with clarity and must reflect the meaning of the scales as developed by the author.

5. Inappropriate norms for computer administered tests (Sampson,1985; Fowler,1985; Brown, 1984; Hofer and Green, 1985). Most tests currently administered by computer were originally developed for a traditional paper and pencil approach. Of the tests previously listed Brown reported that only the Rorschach Interpretive System, Visual Searching Task, and Behavioral Observation Checklist were developed specifically for, and normed on, the computer. Differences in mode of administration may make paper and pencil norms inappropriate for computer administered tests. Hofer and Green reported that many of the computerized tests have not been tested for equivalence to their traditional counterparts and that until such time as equivalence data are available "interpretation of computer-obtained scores with conventionally obtained data should be rejected if there are plausible reasons for expecting nonequivalence"

(p.832).

The literature suggests several major implications for the future development and use of computer-assisted testing. Hofer and Green, (1985) envision that in the decade of the 1990's every psychologist will be utilizing computerized assessments as the principal means for collecting standardized data on clients. Before this occurs Sampson (1984) feels that there is a need for additional research on computer applications in the testing and assessment fields. Sampson has identified the following need areas:

1. There is a need for research on the effectiveness of computer assisted testing and assessment with various client populations.
2. There is a need to assess what human factors affect client use of computer-assisted testing and assessment systems to insure the development of programs which are easy to use and understand.
3. There is a need for research on the effectiveness and validity of generalized computer controlled test interpretations.
4. There is a need for research into the general reliability and validity of computer-assisted tests.
5. There is a need to determine whether existing paper and pencil norms are appropriate for computer-assisted tests.
6. There is a need for the development of initial screening devices to determine whether potential users are psychologically and intellectually capable of using a computer system.
7. Preservice graduate training programs and inservice professional development programs need to include exposure to computer-assisted testing and assessment systems. A final issue which needs to be examined before computer-assisted assessment becomes a commonplace practice is the development of ethical guidelines and standards for the use of computerized tests and test interpretations (Fowler, 1985). Foremost in this area is the question of who is qualified to use computerized tests and computer-based interpretive reports. Fowler reported that at present there are no clear-cut standards to determine who is qualified to use psychological tests and who is not. Hofer and Green(1985) proposed a challenge of competence in computerized psychological testing, which "involves not only

developing standards of practice and responsibility for test developers, publishers, and users but also involves developing a consensus among professionals and getting this consensus formalized in professional standards; in contracts among practitioners, developers, and publishers; and even in state laws governing who is qualified and licensed to give computerized tests" (p.827).

Chapter 3

Methodology

In this chapter the methodology for examining the comparability of the W.A.I.S.-R. and the M.A.B., and answering the questions posed in chapter 1 will be set out.

Subjects

The testing for this study was conducted between May and July, 1986 at the University of Alberta Clinical Services in the Faculty of Education. Subjects were 40 adults who responded to a request for volunteers, broadcast on a local radio phone-in show called "That's Living". (The program, aired Mondays through Fridays, is hosted by two psychologists who discuss topics of psychology and problems of everyday living.) Testing times were arranged with subjects at their convenience. The majority of subjects were given both the W.A.I.S-R. and the M.A.B. in a single three to three and a half hour session. Several subjects were unable to commit to this length of session and were tested on two separate occasions, each of which lasted approximately one and a half hours. Subjects interested in taking the test for career purposes were also given a Jackson Vocational Interest Survey. Two subjects were screened out due to reading disabilities. Subjects were typically individuals who were unemployed or considering career changes. Three of the subjects were currently being treated by a psychiatrist for depression.

Procedure

Subjects were randomly assigned to one of four conditions as they were contacted to arrange appointments for testing. In condition 1 (W.A.I.S-R - M.A.B. Computer), subjects were first administered the W.A.I.S.-R. and then were administered the computerized version of the M.A.B.. In condition 2 (W.A.I.S.-R.-M.A.B. P.and P.), subjects were first administered the W.A.I.S.-R. and then were given the pencil and paper version of the M.A.B.. In condition 3 (M.A.B. Computer

- W.A.I.S.-R.), subjects were first administered the computer version of the M.A.B. and then were administered the W.A.I.S.-R.. In condition 4 (M.A.B. P. and P. - W.A.I.S.-R.), subjects were first administered the paper and pencil version of the M.A.B. and then were administered the W.A.I.S.-R.. As the W.A.I.S.-R. takes from 60 to 90 minutes to administer and the M.A.B. takes another 90 minutes, a short "coffee break" took place between the two testings. Thirty-four of the subjects were seen by the researcher and six were tested by an assistant. Both the researcher and the assistant had formal graduate training in the administration of the W.A.I.S.-R.

Upon completion of the two tests, a sample of the subjects was interviewed to discover what strategies they had used during each of the performance tasks on the W.A.I.S.-R. and the M.A.B.. Subjects were also asked to comment on which test they preferred and what they liked and disliked about the tests. Subjects were asked for permission to record their responses and all post interviews were taped. Selection of subjects for this portion of the study was based on whether or not there was time left at the end of each session.

Prior to testing subjects were requested to read and sign a research release form (Appendix 1) and given a brief description of the study. Background information was also collected from each client for file purposes. All subjects were given verbal and written feedback on the results of their tests with the option of receiving these by phone and mail or in person during a second interview.

Description of Test Instruments

The W.A.I.S.-R is comprised of eleven subtests, six of which constitute the Verbal Scale and five the Performance Scale. These subtests are listed and briefly described below. They are numbered in order of their administration, in which verbal and performance tests are alternated. For Arithmetic, Picture Arrangement, Block Design, Object Assembly, and Digit Symbol, both speed and correctness of performance influence the score. There is no time limit on the other 6 subtests.

Verbal Scale

1. Information: 29 questions covering a wide variety of information that adults have presumably had an opportunity to acquire in our culture. Measures long term memory, education, intellectual curiosity, experiential and reading background, and general Knowledge.

3. Digit Span: Orally presented lists of three to nine digits are to be orally reproduced. In the second part, the client must reproduce lists of two to eight digits backwards. Measures short term auditory memory for unstructured information.

5. Vocabulary: 35 words of increasing difficulty are presented both orally and visually. The examinee is asked what each word means and is scored 2,1, or 0, according to the level of understanding expressed. Measures expressive vocabulary and verbal fluency.

7. Arithmetic: 14 problems similar to those encountered in elementary school arithmetic. Each problem is orally presented and is to be solved without the use of pencil and paper. Measures auditory memory, numerical reasoning abilities, attention, and concentration.

9. Comprehension: 16 items, in each of which the client explains what should be done under certain circumstances, why certain practices are followed and the meaning of proverbs. Each item is scored 2,1, or 0, depending on the degree of understanding expressed and the quality of the response. Designed to measure degree of social acculturation, practical knowledge and common sense.

11. Similarities: 14 items requiring the client to say in what way two things are alike. Each item is scored 2,1, or 0, depending on the degree of understanding expressed and the quality of the response. Indicates level of awareness of relationships noted as concrete, functional or abstract.

Performance Scale

2. Picture Completion: 20 cards, each containing a picture from which some part is missing. The client must tell what is missing from each picture. Measures visual acuity and attention to detail.

4. Picture Arrangement: Each of the ten items consists of a set of cards containing pictures to be rearranged in the proper sequence so as to tell a story.

Measures the ability to analyze, interpret, and predict the outcome of social situations.

6. Block Design: Uses a set of 9 cards containing designs in red and white and a set of identical one inch blocks which are painted red, white, and red and white. The client is shown one design at a time, which he must reproduce using the blocks. Measures the ability to analyze wholes into their component parts, level of nonverbal reasoning, and visual motor coordination.

8. Object Assembly: In each of the four parts of this subtest, cut outs are to be assembled to form a flat picture of a familiar object. Measures the ability to form a coherent whole from parts, the ability to understand the relationship between parts, and visual motor coordination.

10. Digit Symbol: The client is presented a key containing 9 symbols paired with 9 numbers and a series of boxes containing numbers in the top half and blanks in the bottom. Using the key, the client has 90 seconds to fill in the symbols which match each of the numbers. Measures sequencing ability, fine motor coordination, and the ability to learn a new task quickly.

The M.A.B. is comprised of 10 subscales, five of which are Verbal and five Performance. The test has a multiple choice format; clients must choose either A,B,C,D,or E. There is a time limit of seven minutes allowed for each subtest and clients may work only on the specified subtest during that time. All of the Verbal subtests are administered first followed by the Performance subtests. For the paper and pencil version the client is read a standardized set of instructions for each subtest, is directed to look at the example problems, and then is timed for each section. The computerized version of the M.A.B. is completely self administered; all instructions are presented to the client on the computer screen, and each subtest is automatically timed. For the performance subtests the client must refer to the test booklet and then answer on the computer. The verbal subtests are fully contained on the computer. The Verbal and Performance scales are listed and described briefly below in the order in which they are administered.

Verbal Scale

1. Information: 40 questions covering a wide variety of information areas that adults have presumably had exposure to.

2. Comprehension: 28 items, in each of which the client must choose what should be done under certain circumstances, the reasons for certain practices, and the meaning of proverbs.

3. Arithmetic: 26 mathematical problems which the client may solve with the aid of a pencil and paper.

4. Similarities: 34 pairs of words and the client must choose the answer which best describes how the two words, objects, or concepts, are alike.

5. Vocabulary: 46 words and the client must choose the answer which is nearest in meaning to the word given.

Performance Scale

6. Digit Symbol: 35 items comprised of series of 1 to 9 symbols. The client must refer to the key at the top of the page and choose the answer which contains the series of numbers which match the symbols in the key.

7. Picture Completion: 35 pictures in which there is an important part missing. The client must identify the missing part and then choose the answer which contains the first letter of the missing part.

8. Spatial: 50 items. The client must choose the figure on the left which can be made to look like the figure on the right by turning or rotating it, but not flipping it over.

9. Picture Arrangement: 21 sets of cartoon pictures in a mixed up order. The client must choose the arrangement which makes the most sensible story.

10. Object Assembly: 20 items comprised of parts of familiar objects. The client must choose the order, from left to right, in which the parts should be placed to form the object.

Research Questions

In order to answer the questions posed at the end of Chapter 1 the following statistical research questions were proposed:

1a. Are there statistically significant differences between I.Q. scores on the

W.A.I.S.-R. and I.Q. scores on the M.A.B. depending on which test was administered first?

1b. Are there statistically significant differences between subscale scores on the W.A.I.S.-R. and subscale scores on the M.A.B. depending on which test was administered first?

2. Are there statistically significant differences between the I.Q. and subscale scores generated by the paper and pencil M.A.B. and the I.Q. and subscale scores generated by the computer M.A.B.?

3. Are there statistically significant differences between the Verbal, Performance, and Full Scale I.Q. scores on the W.A.I.S.-R. and the Verbal, Performance and Full Scale I.Q. scores on the M.A.B.?

4. Are there statistically significant differences between the subscale scores on the W.A.I.S.-R. and the subscale scores on the M.A.B.?

5. What is the correlation between the W.A.I.S.-R. Verbal, Performance, and Full Scale I.Q. scores and the M.A.B. Verbal, Performance, and Full Scale I.Q. scores respectively?

6. What is the correlation between the W.A.I.S.-R. subtest scores and the M.A.B. subtest scores?

Analysis

In order to answer the above questions the following analyses were used. For questions 1a. and b., Hotelling T^2 tests were performed to determine if there was a practice, or sequence, effect depending on whether the W.A.I.S.-R. was administered first or the M.A.B. was administered first. Question 2, which addressed whether the different forms of the M.A.B. produced equivalent scores, was examined using a Hotelling T^2. For questions 3 and 4 , which addressed the possible differences between the scores on the W.A.I.S.-R. and the scores on the M.A.B., a correlated t test was used. Questions 5 and 6, which examined the relationship between the scores on the W.A.I.S.-R. and the M.A.B., were analyzed using Pearson Product Moment Correlations. The level needed for statistical significance for all of the above questions was alpha = .05. The question of what cognitive strategies individuals used in solving certain subtests and which test

they preferred was examined using a descriptive report format.

In the following chapter the results of the statistical analysis for each of the above proposed questions will be summarized. Chapter 4 will also include a description of the test population, a description of the statistical analysis used, and a summary of the information concerning cognitive strategies which was gathered during post test interviews.

Chapter 4

Results

Description of the Sample

A total of 40 subjects were tested in this study. All 40 subjects were administered the W.A.I.S.-R.; 20 of these received the M.A.B. paper and pencil version and 20 received the M.A.B. computer version. Of the 40 subjects 20 were female and 20 were male. The M.A.B. paper and pencil group was comprised of 12 females and 8 males. The M.A.B. computer group consisted of 8 females and 12 males. Ages for the 40 subjects ranged from 16 to 60 years, with a mean age of 32. Education level in years of schooling ranged from 9 to 18 with a mean number of 13 years. The mean Full Scale I.Q.s for the W.A.I.S.-R. and the M.A.B. were 112 and 108 respectively, which is 12 and 8 points higher than the mean for the general population. On the variables of age, education, and I.Q., the M.A.B. paper and pencil and computer groups were very similar.

Sequence Effect of Test Taking

In order to see if there was a sequence, or practice, effect of taking the M.A.B. either before of after the W.A.I.S.-R., a Hotelling T^2 test was run comparing the scores on the test taken first (W.A.I.S.-R. or M.A.B.) with the test taken second (W.A.I.S.-R. or M.A.B.). Results of the Hotelling T^2 (T^2= 0.263, p= 0.754 for the W.A.I.S.-R. and T^2= 0.372, p=0.407 for the M.A.B.) indicate that there was no significant difference between the group means of the tests taken first or second. Therefore, it was decided to drop any sequencing effect from further analysis.

Comparability of the Different Forms of the M.A.B.

In order to determine if the paper and pencil form and computer form of the M.A.B. produce equivalent scores a Hotelling T^2 was performed comparing the scores on the paper and pencil M.A.B. and the scores on the computer M.A.B. The results (T^2= 0.356, p= 0.442) indicated that there were no significant differences

between the group means of the paper and pencil test and the computer test. Since there were no significant differences between the two forms of the test it was decided to collapse the scores from the paper and pencil and computer test for further analysis. A final Hotelling T^2 was performed to check for sequence-form interaction effects. The results (T^2= 0.344, p= 0.626 for the W.A.I.S.-R. and T^2= 0.116, p= 0.970 for the M.A.B.) indicated that there was no significant interaction between the form of test and the order in which it was given.

Correlated t Test Results

A correlated t test was performed comparing equivalent subtest and I.Q. scores on the W.A.I.S.-R. and the M.A.B. to determine if there were significant differences between the scores on the two tests. The results of the analysis are presented in Table I. Statistically significant mean differences were found on eight of the ten comparable subscales. The largest subscale mean score differences were found on Comprehension, Similarities, Picture Completion, Object Assembly, and Digit Symbol. For the first three aforementioned subtests, mean scores were higher on the W.A.I.S.-R.; for the last two, mean scores were higher on the M.A.B. Statistically significant differences were not found for the Arithmetic and Picture Arrangement subscales. In terms of I.Q. scores, statistically significant differences were found for the Verbal I.Q. and the Full Scale I.Q. but not for the Performance I.Q.. The differences between the means for both the Verbal and Full Scale Scores were 4 I.Q. points, with scores on the W.A.I.S.-R. being higher in both cases.

Correlations

In order to examine the relationship between the scores of the W.A.I.S.-R.

Table 1

Means, Standard Deviations, t-values, and Probabilities of Differences between Means of Subtest Scores and I.Q. Scores on the W.A.I.S.-R and the M.A.B.

Subtest	W.A.I.S.-R. Mean	St. Dev.	M.A.B. Mean	St. Dev.	t-value	p
Information	55.00	8.20	52.47	9.17	2.31	*0.026
Vocabulary	58.05	7.10	55.85	8.63	2.38	*0.023
Arithmetic	54.60	8.92	55.82	7.65	-0.85	0.401
Comprehension	60.77	6.75	54.37	6.13	4.98	*0.000
Similarities	59.52	7.93	56.10	6.74	3.04	*0.004
Picture Completion	54.32	6.70	50.67	7.51	3.05	*0.004
Picture Arrangement	53.85	9.06	55.70	10.86	-1.24	0.221
Block Design/Spatial	54.82	7.92	51.80	10.26	2.37	*0.023
Object Assembly	52.32	7.95	55.75	7.76	-3.44	*0.001
Digit Symbol	49.62	7.50	54.02	8.58	-3.50	*0.001
Verbal I.Q.	112.85	12.41	108.45	10.98	3.39	*0.002
Performance I.Q.	109.27	11.56	109.47	12.45	-0.17	0.864
Full Scale I.Q.	112.45	12.11	108.82	11.41	3.24	*0.002

Degrees of freedom N-1=39

t =2.02; p=0.05

* Standard deviations for subscales are reported in z score points. S.D.s for Verbal, Performance and Full Scale I.Q.s are reported in I.Q. points.

Table II

Correlations Between The W.A.I.S.-R. and The M.A.B.

	W.A.I.S.-R.												
M.A.B.	1	2	3	4	5	6	7	8	9	10	V	P	F
1	*.688	.672	.438	.377	.531	.360	.515	.468	.473	.263	.699	.520	.722
2	.627	*.739	.557	.249	.489	.300	.555	.279	.330	.254	.629	.521	.665
3	.352	.541	*.402	.380	.415	.264	.246	.396	.268	.076	.558	.278	.498
4	.446	.488	.282	.207	.546	.111	.332	.327	.270	.422	.551	.306	.464
5	.408	.537	.222	.290	*.539	.201	.373	.455	.546	.608	.439	.477	.521
6	.398	.457	.344	.217	.309	*.438	.359	.452	.482	.354	.358	.487	.465
7	.285	.476	.323	.250	.228	.349	*.567	.573	.596	.352	.382	.643	.565
8	.393	.562	.310	.289	.217	.418	.448	*.632	.624	.270	.463	.667	.637
9	.157	.280	.187	.259	.226	.131	.293	.673	*.679	.486	.346	.551	.499
10	.265	.224	.167	.161	.202	.236	.354	.604	.495	*.519	.222	.525	.406
V	.624	.705	.439	.373	.574	.276	.480	.492	.467	.303	*.760	.547	.764
P	.424	.592	.440	.283	.308	.405	.535	.753	.731	.414	.517	*.816	.738
F	.569	.716	.508	.348	.487	.372	.550	.676	.663	.398	.699	.744	*.821

* indicates corresponding subtest and I.Q. correlations that are significantly different from zero.

The following key indicates the subtests which correspond to the numbers in table II: 1-Information; 2-Vocabulary; 3-Arithmetic; 4-Comprehension; 5-Similarities; 6-Picture Completion; 7-Picture Arrangement; 8-Block Design/Spatial; 9-Object Assembly; 10-Digit Symbol; V-Verbal I.Q.; P-Perfomance I.Q.; F-Full Scale I.Q..

and the M.A.B., Pearson Product Moment correlations between comparable subtest and I.Q. scores were computed. The results of the analysis are summarized in Table 2 in the form of a correlation matrix. Correlations for the subtests range from a low of 0.20 for Comprehension to a high of 0.74 for Vocabulary. Generally, subscale correlations are in the range of 0.40 to 0.67. Correlations for the Verbal I.Q., Performance I.Q., and Full Scale I.Q. are 0.76, 0.81 and 0.82 respectively.

Cognitive Strategies and Test Preference

Upon completion of the two tests 15 subjects were interviewed to determine whether they used the same cognitive strategies to complete comparable subtests on the W.A.I.S.-R. and the M.A.B.. Of these subjects 7 were female and 8 were male; 11 completed the M.A.B. on the computer and 4 wrote the paper and pencil version. The subtests that were examined were Object Assembly, Picture Arrangement, Digit Symbol, Block Design, and Spatial. These are the subtests which differ most in format on the two tests.

Subjects experienced a great deal of difficulty in describing the strategies they had used. In many cases the descriptions they gave were vague and they could not elaborate when prompted to do so. In addition, subjects often reported using the same strategies on both tests, but the actual strategies which they described differed for each test. For example, on the Spatial and Block Design subtests many subjects reported that they had used the same strategies on both subtests but when asked to describe how they solved each task they gave different a description of strategies used to solve the Spatial subtest than the description they gave for the Block Design subtest. Cognitive strategies and their implications for future research will be further discussed in Chapter 5. The following is a summary of the strategies which subjects reported using.

For the Object Assembly subtest subjects reported using basically the same strategies with two slight differences between those used on the W.A.I.S.-R. and those used on the M.A.B. First, on the W.A.I.S.-R., subjects typically looked for a major identifiable part and then added pieces to it, while on the M.A.B. they typically worked from left to right or found a beginning and an end and then filled

in the center. Secondly, the W.A.I.S.-R. involved physical manipulation of the pieces while the M.A.B. involved mental manipulation. Subjects did however, report visualizing what the object would look like as a whole on both tests. Two of the subjects reported using a strategy on the M.A.B. which cannot be used on the W.A.I.S.-R. This entails deciding on which pieces should come first and last, checking the answer choices given, and eliminating the items which do not have the pieces in that specific order.

For Picture Arrangement, all of the subjects reported using basically the same strategies on both tests, although individual subjects differed as to the strategy they used. The most common strategy used was to find a beginning and an end and then fill in the middle. Three of the subjects also reported using the choice of answers given, on the M.A.B., as additional clues to the correct sequence.

On the Digit Symbol subtest the majority of subjects reported using the same strategy on both tests, of referring to the key until they had most of the symbol/number pairs memorized. A few of the subjects were unable to memorize the symbol/number pairs and had to constantly refer to the key. Again, three subjects reported using the choice of answers given, on the M.A.B., as additional clues.

The Block Design subtest on the W.A.I.S.-R and Spatial subtest on the M.A.B. measure the same cognitive ability according to Jackson(1984). However, because of the widely different formats used on each of the tests the strategies for solving each will be discussed separately and their comparability will be examined in the chapter on discussion and conclusions. For the Block Design subtest subjects used several different strategies. Some subjects mentally divided, or broke the figure up into squares. Some subjects started with a corner and worked across the the figure or started at the top and worked down, while others simply used trial and error. A number of subjects reported focussing on the whole design while others focussed only on certain aspects or the most salient features of the design.

For the Spatial subtest the majority of subjects reported turning each geometric design in their mind until it matched the key design. A large portion of

these subjects focussed on one corner of the design, or an identifying mark such as a dot, as a point of reference, rather than looking at the design as a whole. Two subjects actually rotated the book rather than manipulating the figures in their mind, and one subject simply guessed. Another subject described a rather unique strategy of mentally picking up each figure and placing it on top of the key to see if it fit.

Once subjects finished their description of cognitive strategies they were asked to comment on which test they preferred and why. The results of the survey are as follows. Out of a total of 15 subjects surveyed, 11 preferred the W.A.I.S.-R., while 4 preferred the M.A.B. Five of the 7 females preferred the W.A.I.S.-R. to the M.A.B.. Of these five, 3 were administered the computerized M.A.B. and 2 were administered the paper and pencil version. Subjects gave varied reasons for preferring the W.A.I.S.-R., the most common of which were: 1. they preferred manipulating with their hands rather than with their minds; 2. they felt more relaxed during the W.A.I.S.-R.; 3. they found the computer hard on their eyes; and 4. they enjoyed the human contact during the W.A.I.S.-R. The two subjects who preferred the M.A.B. to the W.A.I.S-R. were both administered the computer test. The reasons given for choosing this test were: 1. they liked the novelty of the computer; 2. they felt anxious with another person present during the W.A.I.S.-R.; 3. they felt embarrassed during the W.A.I.S.-R.; and 4. they liked the anonymity of the computer.

Six of the 8 males surveyed preferred the W.A.I.S.-R. to the M.A.B. Of these, 4 were administered the computerized M.A.B. and 2 were administered the paper and pencil version. The reasons given for their preference were: 1. they liked the personal contact with the W.A.I.S.-R.; 2. they liked the "hands on" format of the W.A.I.S.-R.; 3. they did not quite feel comfortable with the computer; and 4. they found it hard to read the screen. Both of the subjects who preferred the M.A.B. were administered the computer version. They gave the following reasons for their preference: 1. they could work on it alone; 2. they felt more at ease; and 3. they were not embarrassed by having another person present.

It should be noted that subjects were surveyed as to which test they preferred and not on whether they liked or disliked the computer administered test.

Out of the total 20 subjects tested on the computer only one subject indicated that he disliked the computerized test. Generally, comments about the computer by the total group of test subjects were favorable. The majority of subjects had never had any contact with micro-computers previous to testing and most felt that the experience had been a positive one. A few subjects did express anxiety concerning the 7 minute time limit for each subtest on the M.A.B. and were frustrated at not being able to complete all items in each test. These comments however apply equally to the computerized M.A.B. and the paper and pencil M.A.B..

The following chapter will include a summary of the results, a discussion of the research findings, and the conclusions which were drawn from the study. Implications for further research will also be set out.

Chapter 5

Summary, Discussion and Conclusions

Summary

The purpose of this study was to determine to what extent the Multidimensional Aptitude Battery is a suitable alternative to the Wechsler Adult Intelligence Scale-Revised. A total of 40 adult subjects were administered the W.A.I.S.-R. and the M.A.B.. Twenty of the subjects received the computerized M.A.B. and 20 received the paper and pencil version. Upon completion of testing a sample of 15 subjects were interviewed concerning the cognitive strategies they had used on certain subtests and which test, W.A.I.S.-R. or M.A.B., they preferred.

The results of the statistical analysis of the test scores indicated that there was no significant practice effect, or sequence effect, when the W.A.I.S.-R. and M.A.B. were administered back to back. The M.A.B. could therefore, be used as a retest instrument when the results of the W.A.I.S.-R. are deemed to be unreliable or invalid, or an indication of change is desired. There were no significant differences found between the mean scores on the computerized M.A.B. and the paper and pencil M.A.B. suggesting that the two tests yield equivalent results. As such, the normative data for the paper and pencil M.A.B. would be equally applicable to the computerized M.A.B., and renorming of the computerized test not needed. For the subscale scores statistically significant differences were found on 8 of the 10 subtests. Closer examination of these suggests that for most subtests the differences are not clinically significant such that in actual practice differences of this size may have little impact on the interpretation of the test results. The clinical significance of the W.A.I.S.-R. and M.A.B. subscale score differences and their implications in terms of practical application of the M.A.B. will be examined in the Discussion section.

Results of the analysis of the Verbal, Performance, and Full Scale I.Q.

scores suggest that the M.A.B. and W.A.I.S.-R. scores are quite comparable. The 4 point difference between the mean Full Scale I.Q.s for the W.A.I.S.-R. and the M.A.B. is only one third of the standard deviation of the W.A.I.S.-R. and less than two times the standard error of measurement. In terms of practical application a difference of this size is negligible as it will not greatly affect the interpretation of the test score. Correlations for the W.A.I.S.-R. and the M.A.B. Verbal, Performance, and Full Scale I.Q. scores are relatively high (.76, .81 and .82 respectively), indicating that the two tests are quite highly related.

Correlations for the subtests range from .20 to .67, suggesting that certain subtests are more highly related than others. The Comprehension and Arithmetic subtest correlations in particular are low, .21 and .40, respectively. Since it has not yet been established that the M.A.B. and W.A.I.S.-R. subtests are measuring the same cognitive processes, using W.A.I.S.-R. clinical interpretations for the M.A.B. subtests should be done with caution.

The results from the interviews can be summarized in the following way. The majority of subjects reported that the strategies which they used on the W.A.I.S.-R. and the M.A.B. were basically the same, however analysis of the way in which they actually solved each task indicates that there may be some differences on certain subtests. The fact that some subtests are of quite dissimilar format also suggests that different cognitive strategies may be used on the W.A.I.S.-R. and the M.A.B. In particular, the subtests of Block Design/Spatial, Object Assembly, and Digit Symbol may be solved in quite different ways on the W.A.I.S.-R. and the M.A.B.. Since subjects experienced a great deal of difficulty in describing the strategies which they used, the results should be interpreted cautiously. Further research using behavioral observations and more in depth interviews is recommended before a conclusion can be reached.

In terms of test preference, the majority of subjects interviewed preferred the W.A.I.S.-R. to the M.A.B.. Of those interviewed and preferring the M.A.B. all were tested on the computer; none of the subjects preferred the paper and pencil M.A.B. Males and females gave similar reasons for their preferences. In general, those who preferred the W.A.I.S.-R. did so because they enjoyed the human interaction and the"hands on" nature of the test. Subjects preferring the

computerized M.A.B. generally did so because they found it novel, and felt more comfortable with the anonymity of a computer. Overall, the computer M.A.B. was well accepted among all twenty of the subjects tested on it.

Discussion

Practical Implications of Score Differences

Intelligence tests such as the W.A.I.S.-R. have two major applications. First, they are used for determining a person's level of expertise in a number of different cognitive areas and for estimating a person's overall level of intellectual ability. This information is obtained from the subscale and I.Q. scores, with the Full Scale I.Q. score being the predominantly important one. In terms of estimating a persons overall level of intellectual functioning the M.A.B. is a suitable alternative to the W.A.I.S.-R. For mean I.Q. scores there is an approximate 4 I.Q. point difference on the Verbal I.Q. and Full Scale I.Q. for the W.A.I.S.-R. and M.A.B., while there is no difference on the Performance I.Q.. This mean difference of approximately two times the standard error of measurement for both tests means, in practical terms, that occasionally when calculating a person's overall intellectual ability on the M.A.B. the result may be a slight underestimate of their ability in terms of W.A.I.S.-R. scores. For example, a person who received a Full Scale I.Q. of 108 on the M.A.B. and a Full Scale I.Q. of 112 on the W.A.I.S.-R. would be classified as having Average intelligence according to the former test and High Average intelligence according to the latter test. However, another person with a similar 4 point difference who obtained an I.Q. of 110 on the M.A.B. and an I.Q. of 114 on the W.A.I.S.-R. would be classified as having High Average intelligence according to both tests. In summary, a difference as large as 9 I.Q. points, which is more than twice the difference between the W.A.I.S.-R. and M.A.B. mean Full Scale scores will result in either the same classification on both tests or a difference of one classification.

The subscale scores on the W.A.I.S.-R. and the M.A.B. are not as comparable as the full scale scores. However, depending on the way in which the scores are reported, as classifications or as percentiles, the differences can be

interpreted as insignificant or as quite large. The largest mean difference, 6.40 Z points, was found on the Comprehension subtest. Differences for the other nine subtests ranged from 3.02 to 4.40 Z points. When the mean scores for the subtests are converted back to scaled scores having a mean of 10 and a standard deviation of 3, with the exception of Comprehension which shows a two point difference, the scores become either equivalent or show a 1 point difference. In terms of practical application using classifications for interpretation a one point difference is negligible. It will have little impact on the interpretation of subtest scores, which are generally classified into the categories of Very Superior, Superior, High Average, Average, Low Average, Borderline, and Mentally Deficient. For example, if a subject received a Z score of 54 on a W.A.I.S.-R. subtest and a Z score of 50 on an equivalent M.A.B subtest, his or her performance would be interpreted as Average on both tests. If a person obtained Z scores of 61 and 54, which are the mean W.A.I.S.-R. and M.A.B scores for the subtest showing the largest difference, his or her performance would be interpreted as High Average and Average respectively. So it can be seen that in terms of interpreting levels of subtest performance using classifications, the W.A.I.S.-R. and the M.A.B. yield similar results.

When reporting results in terms of percentiles there can be a significant difference between a z score of 54 and a z score of 61. A z score of 54 is equivalent to a percentile of approximately 55, while a z score of 61 is equivalent to a percentile of 85. In this case the M.A.B. and W.A.I.S.-R. clearly do not yield similar results suggesting, that the interpretations for these two scores would be different.

The second major application for the W.A.I.S.-R. is that of diagnostic tool. When using W.A.I.S.-R. profiles for diagnosing learning disabilities, brain damage and psychiatric disorders, we infer that each subtest is tapping a unique cognitive ability, or set of abilities, associated with specific underlying cognitive processes. Research over the years on the W.A.I.S.-R. has resulted in hypothesized cognitive abilities for each of the subtests, which are now widely used for clinically interpreting W.A.I.S.-R. profiles. Before the M.A.B. can be used as a valid instrument for making diagnoses and clinical interpretations, psychologists must

understand what each subtest is measuring. It is difficult, from the subjects descriptions of strategies used during the present study, to make inferences about the cognitive processes tapped by the M.A.B. subtests, or to draw conclusions concerning whether the M.A.B. and the W.A.I.S.-R. are actually measuring the same skills. Correlations between comparable subtests would suggest that the majority of M.A.B. and W.A.I.S.-R. subtests are moderately related. Some subtests however, correlate higher with other subtests than their W.A.I.S.-R. or M.A.B. counterparts. Common sense on examination of the subtest formats suggests that there are major differences in the strategies which must be used to solve comparable subtests. However, differing strategies may still be tapping the same underlying cognitive processes. The results of the present enquiry suggest that using W.A.I.S.-R. clinical interpretations for the M.A.B. should be done with caution. However, it is beyond the scope of this study to reach any firm conclusions as to whether the M.A.B. and W.A.I.S.-R subtests are measuring the same cognitive abilities. It is certainly a question which deserves further enquiry.

Comparison to Related Research

Table 3 displays a comparison of correlation coefficients between the W.A.I.S.-R. and the M.A.B., and the W.A.I.S. and W.A.I.S.-R. from research previous to the present study. Correlations found in the present study are somewhat different from those found in the Hirsch study (1986) which examined the comparability of Verbal Scale scores and Verbal I.Q. scores from a computerized M.A.B. program, developed for the study, to the paper and pencil M.A.B. and the W.A.I.S.-R.. Generally, Hirsch's correlations for the verbal subscales (.55 to .85) are higher than those found in the present study (.21 to .76). The largest difference between the two studies was found on the Comprehension subtest: .21 for the present study as opposed to .68 and .77 for the Hirsch study. In terms of Verbal I.Q. scores the correlations are more comparable: .86 for the Hirsch study and .76 for the present study. Differences in correlations between the two studies may be due to the use of different test samples. The subjects in the Hirsch study were a homogeneous group of adults who were either students or

applicants at the Alberta Vocational Center (A.V.C.). A.V.C. is an assessment and training center for socially, economically, and educationally disadvantaged adults. In terms of education level the the mean grade equivalent was 9.5. and the mean M.A.B Verbal I.Q. of 93 for this group was 7 points below the mean for the general population. In contrast, the subjects for the present study were a diverse group of adults with a mean M.A.B. Verbal I.Q. of 108, which is 8 points above the mean for the general population, and a mean grade equivalent of 13.

Correlations for the Verbal subscales, Performance subscales, and Full Scale I.Qs are comparable to, and in some cases higher than, those reported by Jackson(1984), with the exception of Comprehension. Those differences that do exist between the two studies are possibly due to differences in the samples used, as well as differences in the size of sample used. Correlations from the present study are also comparable to W.A.I.S./W.A.I.S.-R. correlations reported by Jackson(1984), again with the exception of Comprehension. An explanation for

Table 3

Comparison of Correlation Coefficients between the M.A.B. and the W.AI.S.-R. and the W.A.I.S. and W.A.I.S.-R.

	Hirsch 1986 N=40		Jackson 1984 N=145	Ackerman 1986 N=40	Wechsler/Smith 1981 N=72/70
Subtest	MAB/WAIS-R Written	Computer	MAB/WAIS-R	MAB/WAIS-R	WAIS/WAIS-R
Information	.57	.77	.82	.69	.84
Comprehension	.67	.59	.73	.21	.62
Arithmetic	.68	.77	.89	.40	.56
Similarities	.55	.66	.66	.53	.52
Vocabulary	.81	.85	.89	.74	.59
Digit Symbol			.45	.52	.54
Picture Completion			.87	.44	.47
Spatial/Block Design			.44	.63	.80
Picture Arrangement			.86	.57	.33
Object Assembly			.65	.68	.40
Verbal I.Q.	.83	.86	.94	.76	.82
Performance I.Q.			.79	.82	.82
Full Scale I.Q.			.91	.82	.87

Note- The WAIS-WAIS-R correlations are based on the arithmetic mean of the correlations reported by Wechsler(1981) and by Smith(1983) and taken from a table in Jackson(1984).

the low correlation for the Comprehension subtest found in the present study as compared to the correlations found in the Hirsch and Jackson studies has not been found. The original data from the Comprehension subtest was checked twice to insure that it had been correctly entered into the computor for statistical analysis and no error was found. The correlation of .20 for the Comprehension subtest may indicate that the two tests are measuring different cognitive functions however, if this were the case one would expect to find low Comprehension correlations in the Hirsch and Jackson studies. An alternative hypothesis for the low correlation is that the M.A.B. and W.A.I.S.-R. Comprehension subtests differ in level of difficulty. Since the sample of subjects in the present study scored an average of 12 points higher on the W.A.I.S.-R. than the general population it may be possible that many of the subjects reached the ceiling on the W.A.I.S.-R. but did not reach the ceiling on the M.A.B.. If this is true then the low Comprehension correlation may be an artifact of the sample tested and would not occur in a sample more representative of the general population.

Quality of the Computerized M.A.B.

Reliability and validity data are not presently available for the computerized M.A.B.. However, because it is identical to the paper and pencil version in all ways except mode of administration one would expect the computerized version to be as valid and reliable as the paper and pencil form. Data on reliability and validity of the paper and pencil M.A.B. as reported in the M.A.B. administration manual was discussed in the review of the literature and indicates that the M.A.B. is of high quality in terms of test construction.

In terms of practical usage the M.A.B. computerized test is of good quality. Ten to fifteen minutes of time is required to learn how to use it and it could easily be administered by nonprofessionals such as psychological assistants and secretaries. It was well accepted by the subjects, most of whom found it a novel experience. None of the subjects had difficulty in following the directions and only a 5 to 10 minute introduction to the computer was needed before subjects were able to take the test independently. For the performance section, referring to the test booklet and then answering on the computer was easily handled by all

subjects, suggesting that using test booklets along with the computer is an acceptable way to administer tests. A few subjects complained about reflection of overhead lights off the computer screen, making it difficult to read. Care should be taken to insure that computers are placed in such a way as to provide optimal legibility of the print on the screen.

The M.A.B. is the only test of its type which offers such a wide range of test items. As well, it covers a wide range of skills associated with intelligence. The M.A.B is patterned after the W.A.I.S.-R., which is a test that has been vigorously studied and has a long history of use, suggesting the M.A.B. has acceptable validity. In addition, administration on the computer ensures standardized administration which is difficult to achieve with individually administered tests which are dependent on the skill of the examiner.

Conclusions

The results of the study suggest that the M.A.B. is a suitable alternative to the W.A.I.S.-R. when used for determining an individuals overall level of cognitive functioning. Psychologists should keep in mind the average 4 point difference between the W.A.I.S.-R. and the M.A.B. on the Verbal and Full Scale I.Q.s when interpreting test results. In terms of the subscale scores the M.A.B. and W.A.I.S.-R. are not equivalent as indicated by statistically significant differences on 8 of the 10 subscales. Depending on the system used for reporting test results, percentiles or classifications, the differences between subscales can appear to be clinically significant or not significant. As well, correlations between the M.A.B. and W.A.I.S.-R. subtests suggest that using W.A.I.S.-R. interpretations for the M.A.B. subtests should be done with caution when making clinical diagnoses of problems such as learning disabilities and organic disorders.

The M.A.B. may be used for retest purposes as there is no practice effect when the W.A.I.S.-R. is administered first. A time lapse of a day between administration of the W.A.I.S.-R. and the M.A.B. is recommended as clients find back to back administration extremely fatiguing. The computerized version of the M.A.B. has distinct advantages. Foremost of these is the saving in cost and time for psychologists. The computerized test can be administered by

nonprofessionals leaving psychologists free for other important clinical work, and enabling them to do more thorough extensive assessments. It does not require lengthy training to administer and can generate results and interpretive reports immediately, giving psychologists the time to spend with clients which would otherwise be devoted to writing reports. In addition, the computerized M.A.B. has the potential to bring the availability of intellectual assessments more in line with the demand, eliminating lengthy waiting periods for assessments.

Using the test booklet with the computer for the Performance section of the test is a perfectly acceptable way to administer the test and has implications for other types of computer administered tests. Designing computer graphics for psychometric tests such as the M.A.B. is an expensive process. Eliminating this cost by developing programs to be used in conjuction with a test booklet helps to make computer-assisted assessments cost effective and widely affordable.

Suggestions for Further Research

In terms of further research, the present study suggests the following areas for investigation:

1. A larger sample of adult subjects drawn randomly from the general population needs to be studied to determine whether the findings from the present study can be generalized.
2. There is a need for research on the computerized M.A.B. to determine its effectiveness with various specific populations such as psychiatric patients.
3. There is a need for the development of a screening device to determine whether a potential user of the M.A.B. is intellectually and psychologically capable of successfully using the computer.
4. There is a need for further research through factor analysis, canonical correlation, or behavioral observations and in depth interviews, to determine whether the M.A.B. and W.A.I.S.-R. subtests are measuring the same cognitive abilities.

REFERENCES

References

Anastasi, A. (1982). Psychological testing (5th ed.). New York: Macmillan.

Baker, F. B. (1984). Technology and testing: State of the art and trends for the future. Journal of Educational Measurement,21(4), 399-406.

Brown, D. T. (1984). Automated assessment systems in school and clinical psychology: Present status and future directions. School Psychology Review, 13(4), 455-460.

Burk, M. J., & Normand, J. (1986, April). Examinee attitudes toward computer-administered psychological testing: A path analytic approach. Paper presented at the annual meeting of the American Research Association, San Francisco.

Fowler, R. D. (1985). Landmarks in computer-assisted psychological assessment. Journal of Consulting and Clinical Psychology, 53(6), 748-759.

Frank, G. (1983). The Wechsler enterprize: An assessment of the development and use of the Wechsler tests of intelligence. New York: Pergamon Press.

Hirsch, B. P. (1986). Comparison of computerized and standardized versions of the Multidimensional Aptitude Battery and the Wechsler Adult Intelligence Scale-Revised. Unpublished master's thesis, University of Alberta, Edmonton.

Hofer, P. J., & Green, B. F. (1985). The challenge of competence and creativity in computerized psychological testing. Journal of Consulting and Clinical Psychology, 53(6), 826-838.

Jackson, D. N. (1984). Multidimentional aptitude battery manual. London, Ontario: Research Psychologists Press.

Lindemann, J. E., & Matarazzo, J. D. (1984). Intellectual assessment of adults. In G. Goldstein & M. Hersen (Eds.), Handbook of psychological assessment (p. 77). New York: Pergamon Press.

Matarazzo, J. D. (1972). Wechsler's measurement and appraisal of adult intelligence (5th ed.). Baltimore: Williams and Wilkins.

Matarazzo, J. D. (1985) . Review of the Wechsler adult intelligence scale-revised. In J. V. Mitchell, Jr. (Ed.) , The Ninth Mental Measurement Yearbook, Vol. 2 (pp. 1703-1705) . Nebraska: Buros Institute of Mental Measurements .

Matarazzo, J. D. (1986) . Computerized clinical psychological test interpretations: Unvalidated plus all mean and no sigma . American Psychologist, 41(1), 14-24 .

Sampson, J. P. (1983) . Computer-assisted testing and assessment: Current status and implications for the future. Measurement and Evaluation in Guidance, 15(3), 293-299 .

Sampson, J. P. , & Pyle, K. R. (1983) . Ethical issues involved with the use of computer-assisted counseling, testing, and guidance systems. The Personnel and Guidance Journal, 61(5), 283-287 .

Sternberg, R. J. (1986) . Intelligence applied: Understanding and increasing your intellectual skills. New York: Harcourt Brace Javanovich .

Wechsler, D. (1958) . The measurement and appraisal of adult intelligence (4th ed.) . Baltimore: Williams and Wilkins .

Wechsler, D. (1981) . The Wechsler adult intelligence scale-revised manual . New York: Harcourt Brace Janovich .

APPENDIX 1

Consent Release Form

Research Consent Form

The purpose of this project is to assess the comparability of intellectual assessments generated by the individual administration of a standardized intelligence test (The Wechsler Adult Intelligence Scale -Revised) and a new test that reports to measure the same intellectual factors (The Multidimensional Aptitude Battery). Your participation in the study will entail being administered either one or both of the above mentioned tests which will take from one and a half to three hours. The results of the tests will be interpreted to you and all information will be kept confidential.

I _____ consent to participate in the above stated research project and give my permission for data collected during these sessions to be used for research purposes. I understand that all information will be kept confidential and that I may withdraw from the project at any time.

Name _____ Date of Birth _____

Address Occupation _____

_____

_____ Education _____

_____

_____ Phone Number _____

Date _____